DØ Note 5906-CONF

# A new expected upper limit on $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ using 5 fb$^{-1}$ of Run II data

The DØ Collaboration
*URL http://www-d0.fnal.gov*
(Dated: March 12, 2009)

We present a new expected upper limit of the rare decay $B_s^0 \to \mu^+\mu^-$ using approximately 5 fb$^{-1}$ of Run II data collected with the DØ detector at the Tevatron. The resulting expected upper limit is $\mathcal{B}(B_s^0 \to \mu^+\mu^-) < 4.3(5.3) \times 10^{-8}$ at the 90%(95%) C.L.

*Preliminary Results for Winter 2009 Conferences*

## I. INTRODUCTION

The standard model (SM) provides an accurate description of current observation from high energy physics experiments, in particular precision electroweak measurements and flavor physics observables. These experiments put strong constraints on extensions of the SM that have tree-level flavor changing neutral current (FCNC) effects. The FCNC decays can only occur at higher order through electroweak penguin and box diagrams in the SM. The SM expectation for the branching fraction of the FCNC decay $B_s^0 \to \mu^+\mu^-$ [1] is $(3.42 \pm 0.54) \times 10^{-9}$ [2]. The decay amplitude can be enhanced by several orders of magnitude in some supersymmetric models through the mediation of the neutral Higgs boson if the value of tan$\beta$ is large. Current observation of this decay at the Tevatron would necessarily imply new physics since the predicted rate for this process in the SM is beyond the detectors' sensitivity. Presently, the best published experimental bound for the branching fraction is $\mathcal{B}(B_s^0 \to \mu^+\mu^-) < 4.7(5.8) \times 10^{-8}$ at the 90%(95%) C.L. [3].

## II. DETECTOR AND DATA SAMPLE

The DØ detector is described elsewhere [4]. The main elements, relevant for this analysis, are the central tracking and muon detector system. The central tracking system consists of a silicon microstrip tracker (SMT) and a central fiber tracker (CFT), both located within a 2 T superconducting solenoidal magnet. The muon detector located outside the calorimeter consists of a layer of tracking detectors and scintillation trigger counters in front of 1.8 T toroidal magnets, followed by two more similar layers after the toroids, allowing for efficient detection out to pseudorapidity $\eta$ [5] of about 2.0. During the spring of 2006 the SMT detector was upgraded by inserting an additional layer of silicon microstrip detectors close to the beam pipe. Data taken before the summer of 2006 are referred to as Run IIa and the data taken afterwards are called Run IIb data. In Run IIb we also have tighter trigger requirements to deal with the increased instantaneous luminosity.

The data used in this analysis is the complete data sample up to December 20, 2008. The integrated luminosity for this sample is roughly 5 fb$^{-1}$. We handle Run IIa and Run IIb data separately; furthermore Run IIb data is split into two subsamples (Run IIb-I, Run IIb-II) based on when the data were recorded. Roughly the integrated luminosity of Run IIa data is 1.3 fb$^{-1}$, Run IIb-I is 1.9 fb$^{-1}$ and Run IIb-II is 1.6 fb$^{-1}$. The main difference between Run IIb-I and Run IIb-II is that there was an upgrade of our trigger table in March 2008 to cope with higher instantaneous luminosity. Therefore data were taken with much higher instantaneous luminosity during the Run IIb-II period. The three data subsamples are treated as three different and independent analyses, but the final upper limit is combined from the separate analyses. To simulate a signal, we have generated Monte Carlo events using the Pythia [6] event generator interfaced with the Evtgen [7] decay package, followed by full Geant v3.15 [8] modeling of the detector response. The Monte Carlo samples for the Run IIa and Run IIb detector configurations have been generated separately.

## III. EVENT SELECTION

The branching fraction of $B_s^0 \to \mu^+\mu^-$ is calculated by

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = N(B_s^0 \to \mu^+\mu^-) \times k, \tag{1}$$

where

$$k = \frac{1}{N(B^+)} \times \frac{\epsilon(B^+)}{\epsilon(B_s^0)} \times f\left(\frac{b \to B^+}{b \to B_s^0}\right) \times \mathcal{B}(B^+ \to J/\psi K^+, J/\psi \to \mu^+\mu^-). \tag{2}$$

$k$ is called the single event sensitivity (*ses*). $N(B_s^0 \to \mu^+\mu^-)$ is the number of signal events, $N(B^+)$ is the number of $B^+$ events in data which is to be used as a normalization in this analysis, $\frac{\epsilon(B^+)}{\epsilon(B_s^0)}$ is the efficiency ratio of $B_s^0$ signal and $B^+$ normalization channel, $f\left(\frac{b\to B^+}{b\to B_s^0}\right)$ is the fragmentation ratio between the $B_s^0$ and $B^+$ events taken to be $3.86 \pm 0.59$, $\mathcal{B}(B^+ \to J/\psi K^+, J/\psi \to \mu^+\mu^-)$ is the branching fraction of the normalization channel obtained $(5.94 \pm 0.21) \times 10^{-5}$. These numbers have been obtained from the 2006 PDG [9] for consistency with the recent CDF result [3].

We only use the dimuon trigger data for this analysis. Figure 1 shows the dimuon invariant mass distribution in data before any kinematics cuts. $B_s^0 \to \mu^+\mu^-$ candidates are formed from pairs of oppositely charged muons. Each muon is required to have a transverse momentum $p_T > 2$ GeV/$c$, and pseudorapidity $|\eta| < 2.0$ to be well inside the

fiducial tracking and muon regions. Each track is required to have at least 3 SMT hits and 14 CFT hits to satisfy a good quality track. To reduce prompt background each muon track is required to be displaced from the primary vertex with $|IP(\mu)|/\sigma_{IP} > 2$, where $IP$ is the impact parameter with respect to the primary vertex. The $B_s^0$ candidate is required to have a good quality decay vertex, displaced from the primary vertex with $L_{xy}/\sigma_{Lxy} > 3$ [10] and the $p_T$ of the $B_s^0$ is required to be greater than 5 GeV/$c$ to reduce dimuon background that is attributed primarily to muons from two separate $B$ decays. A sample of $B^+ \to J/\psi K^+$ events is collected to serve as a normalization channel using the same baseline requirements, but including a requirement of $p_T > 1$ GeV/$c$ for the kaon candidate. The offline reconstruction efficiency between signal and normalization channel largely cancels in the ratio with the exception of the kaon efficiency from the $B^+$ decay. The efficiency ratio is evaluated using Monte Carlo simulation.

To further reduce the background, we construct a boosted decision tree (BDT) [11] classifier that uses five input variables. The $B_s^0$ variables used in the BDT are: Isolation [12] (Fig. 2), Transverse momentum (Fig. 3), Transverse decay length significance (Fig. 4), Impact parameter significance (Fig. 5) and logarithm of vertex $\chi^2$ probability (Fig. 6). To train the BDT, we use signal Monte Carlo events as the signal sample and data sideband events for the background sample, where the sideband is defined as $M_B(5.3663\ \mathrm{GeV}/c^2)$ [13] from $\pm$(3 to 8) $\sigma(M_{Bs})$, where $\sigma = 0.115\ \mathrm{GeV}/c^2$. The BDT has been trained for each of the three data sets separately. Figure 7 shows the distributions of the BDT output for signal and background. To find an optimal BDT cut position, we maximize $\mathcal{P}$ proposed by G. Punzi [14], $\mathcal{P} = \mathrm{S}/(\alpha/2 + \sqrt{\mathrm{B}})$, where S is the number of signal events and B is the number of background events. The constant $\alpha$ is the number of sigmas corresponding to the confidence level at which the signal hypothesis is tested and has been set equal to 2, corresponding to a 95% C.L. After scanning for maximal $\mathcal{P}$, we have found an optimal cut position of the BDT at 0.532 for Run IIa data, 0.608 for Run IIb-I data and 0.631 for Run IIb-II data. Figure 8 shows the dimuon invariant mass distributions after applying the BDT cut. We apply the same BDT cut on the normalization channel, $B^+$, and find $1847 \pm 49(stat.) \pm 115(syst.)$ $B^+$ signal events in Run IIa data, $2188 \pm 52(stat.) \pm 123(syst.)$ events in Run IIb-I data and $1683 \pm 46(stat.) \pm 112(syst.)$ events in Run IIb-II data, where the systematic uncertainty is coming from parametrization of the $B^+$ mass shape. Figure 9 shows the $J/\psi K^+$ invariant mass distributions after applying the BDT cut for each data set. The resulting single event sensitivities with all the relative uncertainties are given in Table I. Uncertainty of the $B_s^0$ momentum is estimated from the difference between Monte Carlo corrections obtained from $B^+ \to J/\psi K^+$ and $B_s^0 \to J/\psi\phi$. Efficiency differences in silicon hit requirement on the kaon track between Monte Carlo and data have been corrected and we assign the statistical uncertainty as one of systematic uncertainties. Aside from the background uncertainty, the largest uncertainty common to the three data sets, 15.2%, comes from the fragmentation ratio $\frac{b\to B^+}{b\to B_s^0}$. The expected SM yields of $B_s^0 \to \mu^+\mu^-$ events in Run IIa, Run IIb-I, and Run IIb-II data sets are $0.192 \pm 0.034$, $0.193 \pm 0.034$ and $0.139 \pm 0.025$, respectively.

## IV. BACKGROUND

The dimuon background in the signal region has been estimated from a fit to the dimuon mass distribution while leaving out the blinded events in the $\pm 3\ \sigma(M_{Bs})$ region around the $B_s^0$ mass. We define the signal region to be $\pm 2.5\ \sigma(M_{Bs})$ around the $B_s^0$ mass. The background shape is parametrized as an exponential plus a flat function, and the fit is performed within the dimuon mass range 4 to 7 GeV/$c^2$. The resulting expected dimuon background events in the signal region are found to be $1.99 \pm 0.62$ for Run IIa data, $3.56 \pm 1.07$ for Run IIb-I data and $2.03 \pm 0.62$ for Run IIb-II data, where the uncertainties are all statistical. In addition we have considered background contributions from $B^0$ and $B_s^0$ decays $B \to h^+h^-$, where $h^\pm$ is a charged kaon or pion. The expected background is calculated based on Eq. 1 and Eq. 2, with $B_s^0 \to \mu^+\mu^-$ replaced by $B \to h^+h^-$. Muon-faking rates for kaon, pion have been estimated using $D^0 \to K\pi$ in $B \to \mu\nu D^0$ decays. The rates are found to be $(0.591 \pm 0.073)\%$ for kaons and $(0.092 \pm 0.059)\%$ for pions, while the muon efficiency is $(72.12 \pm 0.52)\%$ estimated using $J/\psi \to \mu^+\mu^-$ events. Based on these fake rates and branching fractions of $B \to h^+h^-$ in Ref. [13], we have estimated possible non-negligible contributions of misidentified $B_s^0 \to K^+K^-$ and $B^0 \to K^+\pi^-$. Table II lists the background contributions from $B_s^0 \to K^+K^-$ and $B^0 \to K^+\pi^-$ decays. The uncertainties on the $B \to h^+h^-$ estimates are dominated by the fake rate and branching fraction uncertainties.

## V. EXPECTED UPPER LIMIT

Assuming no signal counts (background only) in the signal region, we compute an expected upper limit on the branching fraction at the 90%(95%) C.L. The number of observed events has been set to the number of background events, 2 events for Run IIa, 4 events for Run IIb-I and 2 events for Run IIb-II. We have used a well documented prescription [15] to compute the upper limit. In this calculation, it is assumed that there are no contributions from

$B^0 \to \mu^+\mu^-$ decays, where the decay is suppressed by $|V_{td}/V_{ts}|^2 \approx 0.04$. The expected 90%(95%) upper limits for the branching fraction are found to be $7.6(9.4)\times10^{-8}$ for Run IIa data, $9.9(11)\times10^{-8}$ for Run IIb-I data and $10(13)\times10^{-8}$ for Run IIb-II data. Taking into account the correlated uncertainties for each of data sets, we derive a combined upper limit for all Run II data. The combined upper limit is then $4.3(5.3)\times10^{-8}$ at the 90%(95%) C.L.

## VI. CONCLUSION

With 5 fb$^{-1}$ of Run II dimuon trigger data collected by the DØ experiment, we have studied the sensitivity to the branching fraction of $B_s^0 \to \mu^+\mu^-$ decays. A new expected upper limit on the branching fraction is $\mathcal{B}(B_s^0 \to \mu^+\mu^-) < 4.3(5.3)\times10^{-8}$ at the 90%(95%) C.L. This has similar sensitivity to the best published upper limit from CDF [3], and improves the previous DØ result [16] by a factor of two. Work to understand and reduce the background as well as to include single muon trigger data is ongoing.

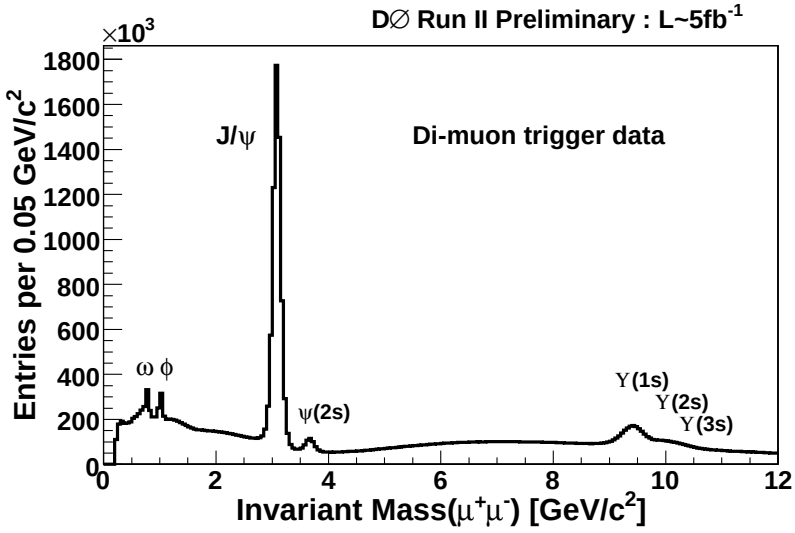


FIG. 1: Uncorrected dimuon invariant mass distribution in the dimuon trigger data set.

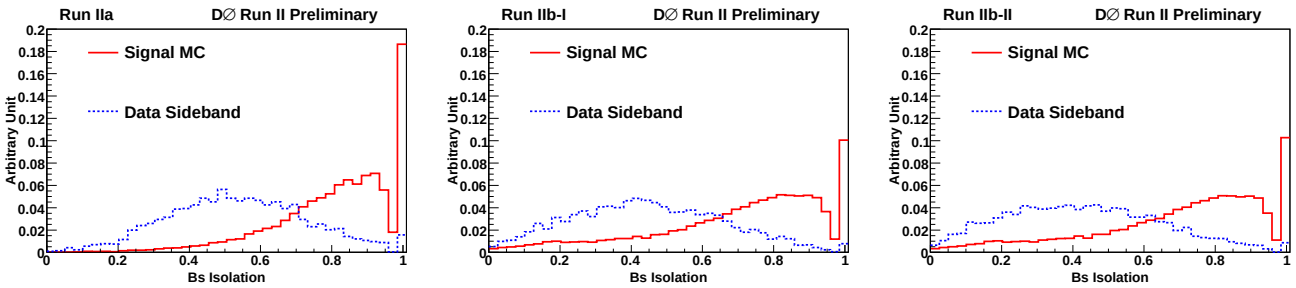


FIG. 2: BDT input, $Isolation(B_s^0)$, for signal and background.

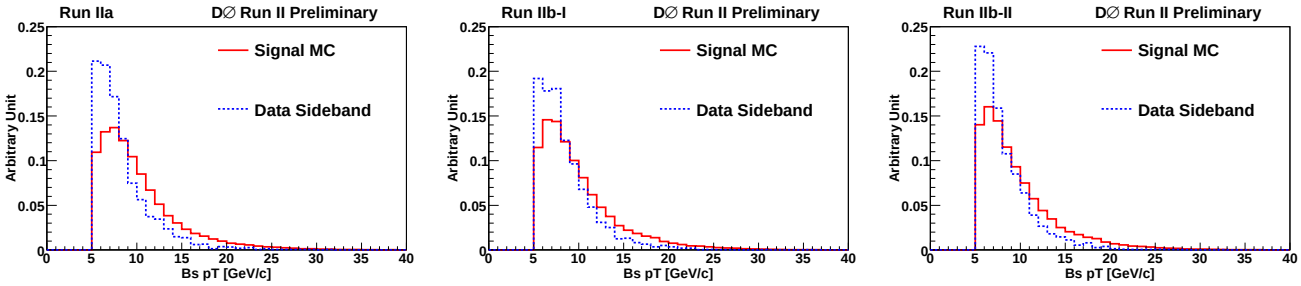


FIG. 3: BDT input, $p_T(B_s^0)$, for signal and background

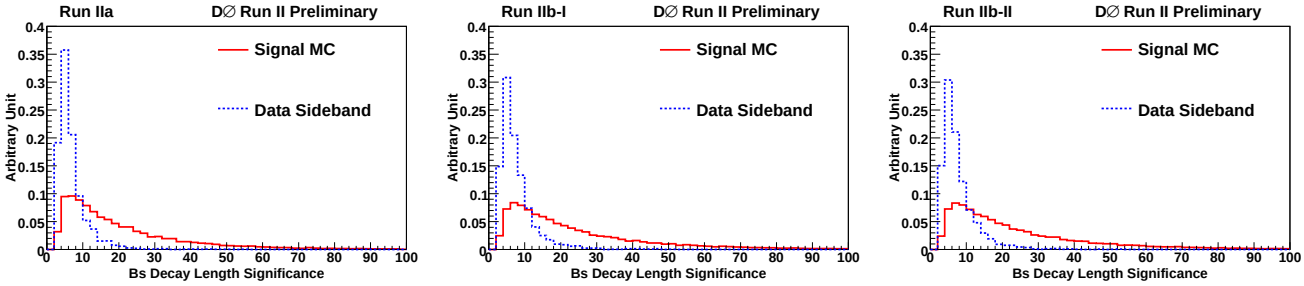


FIG. 4: BDT input, $L_{xy}(B_s^0)/\sigma_{Lxy}$, for signal and background.

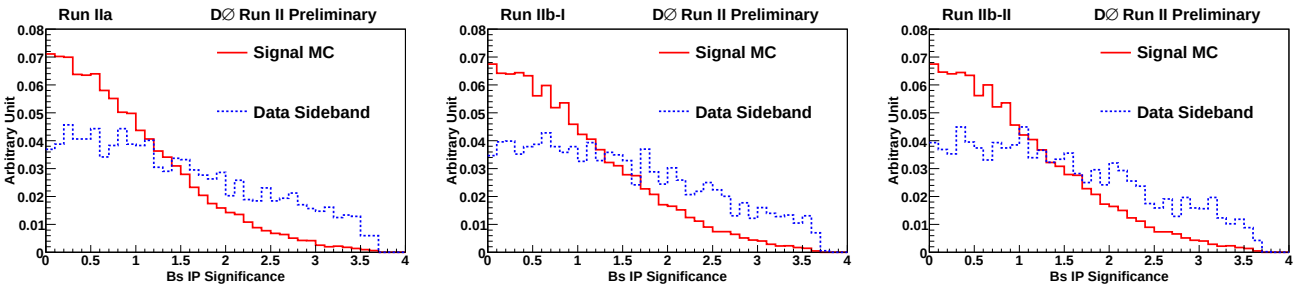


FIG. 5: BDT input, $|IP(B_s^0)|/\sigma_{IP}$, for signal and background.

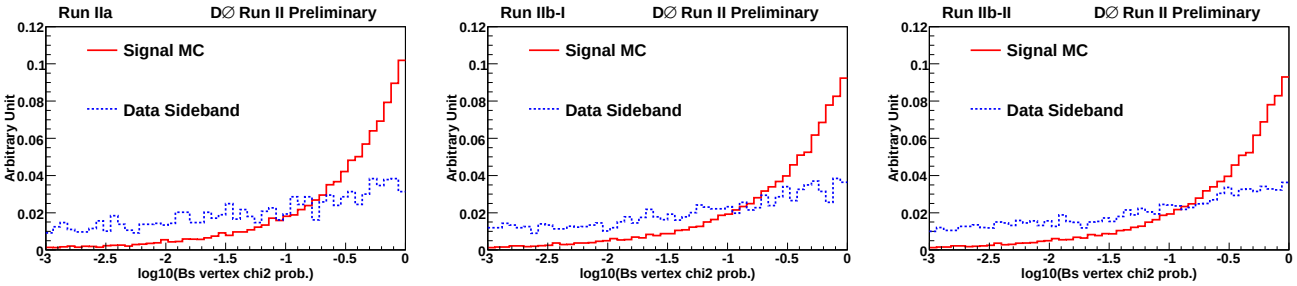


FIG. 6: BDT input, $log_{10}(B_s^0$ vertex $\chi^2$ prob.), for signal and background.

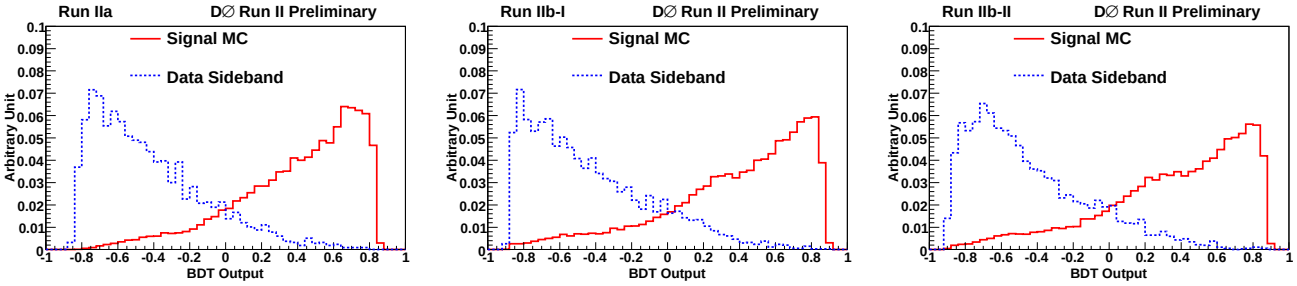


FIG. 7: Distributions of the BDT output for simulated $B_s^0 \to \mu^+\mu^-$ signal and observed sideband events.

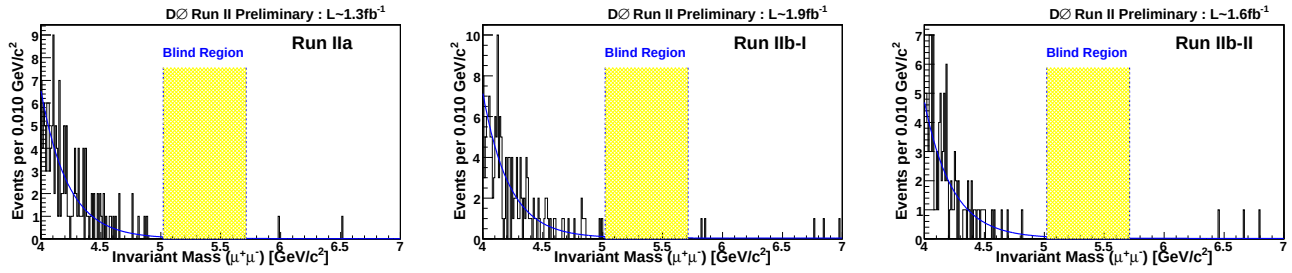


FIG. 8: Dimuon invariant mass distributions after the BDT cut. Search box remains blinded.

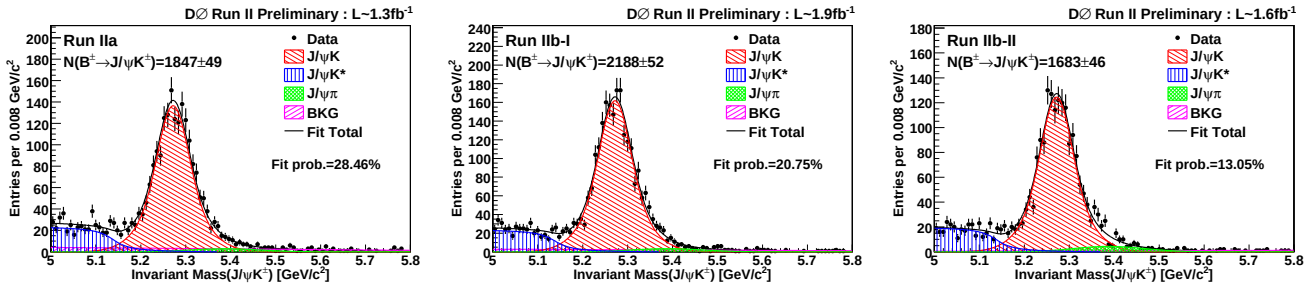


FIG. 9: $J/\psi K^+$ events in data after applying the BDT cut for each data set.

TABLE I: Single event sensitivity ($ses$) and relative uncertainties(%).

| - | Run IIa | Run IIb-I | Run IIb-II |
|---|---|---|---|
| $ses$ | $(1.78 \pm 0.32) \times 10^{-8}$ | $(1.77 \pm 0.31) \times 10^{-8}$ | $(2.46 \pm 0.45) \times 10^{-8}$ |
| $N(B^+)$ stat. | 2.7 | 2.4 | 2.7 |
| $N(B^+)$ syst. | 6.2 | 5.6 | 6.6 |
| # of Si hits on $K^+$ | 0.5 | 0.4 | 0.4 |
| Trigger | 0.3 | 0.6 | 3.4 |
| MC stat. | 0.3 | 0.3 | 0.4 |
| $p_T(B_s^0)$ | 4.4 | 4.5 | 4.5 |
| $f(\frac{b\to B_s^0}{b\to B^+})$ | 15.2 | 15.2 | 15.2 |
| $\mathcal{B}(B^+ \to J/\psi K^+, J/\psi \to \mu^+\mu^-)$ | 3.4 | 3.4 | 3.4 |
| Total | 17.7 | 17.5 | 18.2 |

TABLE II: Number of background events.

| - | Run IIa | Run IIb-I | Run IIb-II |
|---|---|---|---|
| $N(B^0 \to K^+\pi^-)$ | $0.044 \pm 0.029$ | $0.045 \pm 0.030$ | $0.032 \pm 0.021$ |
| $N(B_s^0 \to K^+K^-)$ | $0.124 \pm 0.051$ | $0.125 \pm 0.051$ | $0.090 \pm 0.037$ |
| Dimuon background | $1.99 \pm 0.62$ | $3.56 \pm 1.07$ | $2.03 \pm 0.62$ |
| Total background | $2.16 \pm 0.62$ | $3.73 \pm 1.07$ | $2.15 \pm 0.63$ |
| SM $N(B_s^0 \to \mu^+\mu^-)$ | $0.192 \pm 0.034$ | $0.193 \pm 0.034$ | $0.139 \pm 0.025$ |

**Acknowledgments**

We thank the staffs at Fermilab and collaborating institutions, and acknowledge support from the Department of Energy and National Science Foundation (USA), Commissariat à l'Energie Atomique and CNRS/Institut National de Physique Nucléaire et de Physique des Particules (France), Ministry of Education and Science, Agency for Atomic Energy and RF President Grants Program (Russia), CAPES, CNPq, FAPERJ, FAPESP and FUNDUNESP (Brazil), Deptertments of Atomic Energy and Science and Technology (India), Colciencias (Colombia), CONACyT (Mexico), KRF (Korea), CONICET and UBACyT (Argentina), The Foundation for Fundamental Research on Matter (The Netherlands), PPARC (United Kingdom), Ministry of Education (Czech Republic), Natural Sciences and Engineering Research Council and WestGrid Project (Canada), BMBF and DFG (Germany), A.P. Sloan Foundation, Research Corporation, Texas Advanced Research Program, and the Alexander von Humboldt Foundation.

---

[1] Charge conjugate states are included implicitly throughout this note.
[2] A. J. Buras, Phys. Lett. B **566**, 115 (2003).
[3] T. Aaltonen, *et al.* CDF Collaboration, Phys. Rev. Lett. **100**, 101802 (2008).
[4] V. M. Abazov *et al.* DØ Collaboration, Nucl. Instrum. Meth. A **565** 463 (2006).
[5] $\eta = -\ln[\tan(\theta/2)]$, where $\theta$ is the polar angle with respect to the beam-line.
[6] T. Sjöstrand and M. Bengtsson, Comput. Phys. Comm. **43**, 367 (1987).
[7] D. J. Lange, Nucl. Instrum. Methods A **462**, 152 (2001).
[8] R. Brun and F. Carminati, CERN Program Library Long Writeup W5013 (unpublished).
[9] W.M. Yao *et el.*, J. Phys. G **33**, 1 (2006).
[10] $L_{xy}$ is defined as the projection of the decay length vector($\vec{l}_{xy}$) on the transverse momentum of the $B$-meson($\vec{p}_T^B$) : $L_{xy} = \vec{l}_{xy} \cdot \vec{p}_T^B / |\vec{p}_T^B|$.
[11] A. Hocker *et al.* TMVA - Toolkit for Multivariate Data Analysis, arXiv:physics/0703039
[12] Isolation $I = p_T^B/(p_T^B + \sum p_T)$ where $p_T^B$ is the transverse momentum of the $B$ meson, and $\sum p_T$ is the scalar sum of the transverse momenta of all other tracks within a cone of $\Delta R < 1$ around the $B$ meson, where $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$ and $\phi$ is the azimuthal angle.
[13] C. Amsler *et el.*, Phys Lett. **B667**, 1 (2008).
[14] G. Punzi, Proceedings of PhysStat2003, (SLAC, 2003), eConf C030908, MODT002 (2003).
[15] T. Hebbeker, L3 Note 2633, (2001).
[16] DØ conf-note 5344.